
Book Section:
Stillwell, JCH orcid.org/0000-0002-8803-2001, Bell, M and Shuttleworth, I (2017) Studying Internal Migration in a Cross-National Context. In: Champion, T, Cooke, T and Shuttleworth, I, (eds.) Internal Migration in the Developed World: Are We Becoming Less Mobile? International Population Studies . Routledge . ISBN 9781472478061

Chapter 3

Studying Internal Migration in a Cross-national Context

John Stillwell, Martin Bell and Ian Shuttleworth

This book draws together empirical material on temporal trends in internal migration in selected countries across the developed world in order to explore whether the decline in migration intensity observed in the USA is evident in other developed societies. If the trends in the USA are observed across a suite of comparator nations, then it becomes plausible to contend that the structural economic and social changes that have taken place across the advanced nations of Europe, North America, Asia and Australasia have acted to reduce the propensity for internal migration. If, on the other hand, each country has experienced different, and possibly unique, temporal trends in migration rates in recent decades, the opportunity for grand theory formulation becomes less attractive. Alternatively, it may be that declining migration intensity is not confined to countries that are economically advanced, but that it is more widely spread across nations at earlier stages of development, as indeed seems to be the case with the evidence presented in Chapter 1, which suggests that there is a general period effect that is acting on all countries to a greater or lesser extent.

The question which is the focus of the book clearly requires a cross-national comparative perspective. Given the weight of research that has focussed on internal migration in all its guises across several disciplines, it is tempting to assume that cross-national analysis is straightforward and analyses of migration behaviour in different countries abound. This, however, is not the case; our basic question is easy to pose but much harder to answer as countries – even those with well-established population data systems – differ in the way that data are collected, in their definitions of migration, and in the spatio-temporal coverage of the data that are available. Consequently, previous cross-national studies of internal migration have tended to focus on comparisons between a relatively small number of selected countries rather than confront the challenges of data collection and harmonisation associated with a more comprehensive set of countries. This book falls primarily into the former category by asking individual experts to produce case studies for seven countries that answer the same general questions without being rigidly prescriptive about the means of doing this. In this sense, it differs from studies designed to compare migration using a standard set of migration indicators in two countries (e.g., Bell et al., 2002) or to compare one dimension of migration amongst a relatively small set of countries (e.g., Long et al., 1988). However, it does also

embrace the work of a research project that involved comparative analysis using consistent indicators of migration intensity, impact and distance for a much larger number of countries across the world – the IMAGE project (Bell et al., 2014, 2015c; see also Chapter 4 of this book).

Whilst the statistical issues associated with data harmonisation are particularly important when it comes to making consistent comparisons between countries, as in the IMAGE project, it is also essential to discuss these issues to better inform the limitations and possibilities of the empirical approaches in the country case study chapters that form the core of this book. There are also other conceptual and methodological issues that underlie the comparative focus of the book. Implicit in the structure of the volume are assumptions about the desirability and feasibility of cross-national comparisons, the focus on nation states as the units of analysis, and the importance of internal migration versus other types of population mobility. In assessing the material that is presented in the country case studies, it is therefore also important to explore these issues further so as to understand the limits (and benefits) of our comparative approach.

The purpose of this chapter is to examine these issues in more detail and thereby help readers interpret the statistical material and the arguments that are presented in subsequent chapters. It begins by discussing some of the benefits and the problems of the comparative case study approach in social science, before going on to justify the focus on selected countries (rather than other spatial units) and the migration of people within them. The section that follows thereafter presents an account of the statistical problems of conceptualising and measuring migration in a way that is comparable between countries, drawing primarily on the work of the IMAGE project but also on the experience of preparing the country case studies for Chapters 5-11. It includes a discussion of the challenges faced not only in measuring time series internal migration but also in undertaking quantitative analyses with the data that are available.

The Comparative Approach

As mentioned in Chapter 1, the comparative approach that has been adopted in this book is often used in social science and has been commonly (although not exclusively) applied by

political scientists. It is a useful approach to generate and to explore hypotheses (Collier, 1993) and can assist in the analysis of similarities and differences between societies. It is usually based on the systematic analysis and discussion of a small number of cases and it is therefore suitable when it is not feasible to work with a large number of observations or when it is impossible to design and conduct an experiment, something which is often precluded by the open-ended and complex nature of many social science questions. A comparative approach can throw light on how differences and similarities occur, whether they are between groups of individuals or geographical areas, but important for its success are the units selected for analysis. Lipjhart (1971) argues for analytical units that have as much as possible in common – otherwise they would not be comparable – but which differ on some key dimensions that are the focus of exploration. The strengths of comparative approaches are that they are suitable for answering many social science questions and they are relatively easy to implement, but their weaknesses lie in the many variables and characteristics which it is impossible to make controls for. There is thus sometimes a danger of overplaying the exceptional and the unique at the expense of the general, but these tensions mean that the method is suitable to tease out how general structural forces are mediated and modified by national circumstances. This type of approach is thus well suited to cope with the issues that are the concern of the book.

A comparison between a selection of more advanced countries permits assessment of the extent to which declines in internal migration are common across the developed world, thereby allowing an assessment of the argument that common forces are operating across advanced societies. However, although this approach can be used to challenge and invalidate the central hypothesis of the book, it cannot be used to prove that it is true. In fact, we are not in the business of accepting and discarding hypotheses in these statistical scientific terms but instead operate on the border between quantitative and qualitative approaches, making informed judgements on the basis of the evidence presented in each of the country case study chapters and attempting to draw conclusions from them.

Plausible arguments might be made that challenge this focus on countries (or states) and also on migration flows within them at the expense of other types of mobility and other units used for analysis. It is, for example, entirely feasible to conceptualise migration along a geographical continuum from the very local to the international/global, differentiated only by the prime motives for moving and with a distinction often drawn for intra-national movement

between shorter-distance residential mobility and longer-distance internal migration (see, for instance, Niedomysl, 2011; Niedomysl et al., 2017). Likewise, the same type of argument can apply to the temporal dimension of migration with long-distance commuting shading into long-term business trips and other types of transience which involve moves of varying degrees of permanence over differing spatial scales (see, for instance, Bell and Ward, 2000). One classic example of a population sub-group whose status as internal migrants is questionable is higher education (HE) students in the UK, living some of the year at an address in the vicinity of their chosen university or HE college but returning on regular intervals to inhabit the parental domicile (Duke-Williams, 2009). The picture is made more complex by the fact that different forms of mobility are often closely connected. International moves are often preceded (or followed) by migration within a country, and international migration can influence internal migration by placing pressure on labour and housing markets. In a similar way, temporary forms of movement such as long-distance commuting or seasonal migration, may substitute for, or morph into, more permanent changes of residence (Bell and Ward, 2000).

Despite these complications, internal migration remains important in its own right. It is one of the prime mechanisms by which labour demand and supply are matched at local and regional spatial scales, and where someone lives is an important determinant of their life chances, educational success and life expectancy. However, the distinction between internal and international migration can sometimes be blurred by the regulations operated by national governments or supra-national organisations to manage migrant flows as well as by motives: job-related reasons, for instance, are important for both international and longer-distance internal migrants. Whereas international migration frequently requires some form of documentation or permission (commonly a passport and often a visa), there are situations where this is not a formal requirement – such as migration between the countries of the UK – or where freedom of movement across borders is accepted − such as movement between those countries in the European Union (EU) Schengen (passport-free) zone. The role of the nation state in protecting its borders has therefore been diminished by supra-national organisations such as the EU and the United Nations (UN) in the interests of allowing individuals freedom of movement for economic reasons or to seek sanctuary as refugees or asylum seekers. However, in recent years, some parts of the world have witnessed a backlash against this sort of arrangement, as evidenced in the reaction of some EU member states to the mass refugee migration from the Middle East and Africa, which many commentators

suggest underpinned the Brexit referendum vote in the UK. Increasingly, nations states are responding to the erosion of control of their political territories by globalisation and supranational political organisations by establishing or strengthening border controls as well as reclaiming sovereignty in areas such as labour law, the regulation of the housing market, welfare provision, and education. These are all arenas which influence internal migration. Moreover, despite the extensive attention given to international migration in the media and in public debate, migrations between regions within countries outnumber movements between countries by a factor of four to one (Bell and Charles-Edwards, 2013), and the difference would be even greater if local moves within regions (residential mobility) were taken into account. The scale of this movement alone, coupled with the changing international context therefore readily justify the national focus adopted in this volume – internal migration still matters.

This section has provided a justification for the general approach to the subject matter of the book. However, it is one thing to set out this framework for the comparative study of internal migration but quite another to measure it in ways that permit cross-national comparisons. The latter is by no means easy to achieve. Indeed, while some statistical or methodological difficulties can be overcome or lessened, others are more intractable. The remainder of this chapter is therefore devoted to examining these statistical and technical issues in more detail, so as to make readers more critically aware when approaching the material presented in subsequent chapters. Insights into these measurement and analysis problems refer particularly to UK, US and Swedish data, but the chapter also draws on the important lessons learnt through the IMAGE project (see Bell et al., 2014; 2015a; 2015b; 2015c; Rees et al., 2016; Stillwell et al., 2016).

Studying Internal Migration in an International Comparative Framework

The chief obstacle to achieving a fully comparable analysis is the difference between national statistical systems for the provision of population data. In the UK, for example, data on internal migration can be obtained from the population census administered by three separate national statistical agencies and from administrative sources such as the National Health Service Central Register (NHSCR) and the Patient Register Data Service (PRDS); in Sweden from a population register administered by the Swedish Tax Agency; and in the United States

from a decennial population census and from surveys such as the American Community Survey (ACS) and the Current Population Survey (CPS) conducted by the US Census Bureau. There is considerable diversity, therefore, between just three of our case study countries, never mind the entire sample.

The types of issues that cross-national comparison raises for migration researchers are categorised into four general groups of problems by Bell et al. (2002): (i) temporal comparability – the interval over which migration is measured, commonly one or five years, but also the availability of comparable time-series data; (ii) the way in which migration is measured – for instance as a transition or an event; (iii) data coverage and quality – for instance, certain population groups can be undercounted or omitted entirely from some collections; and (iv) differences in spatial units – this is the well-known modifiable areal unit problem (MAUP) which makes it difficult to compare meaningfully countries with different administrative geographies. An example of the latter is a comparison of inter-state migration in the USA with inter-regional migration in the UK. These are very different geographies at different spatial scales, and hence the type of migration and the distance of move might also be very different. These problems make comparative measurement and analysis a difficult exercise and are discussed in turn below. Some of the problems are important for comparisons between places and through time, whereas some are more important for one or the other. Certain limitations can be addressed and their effects ameliorated whereas others are more intractable and, in making comparisons through time and between places, our only solution is to be aware of the potential difficulties and to make an informed judgement about the evidence that is presented.

Measuring Migration

One important difference between population data systems is the way in which they record and conceptualise migration. Population registers and administrative data systems usually record migration as an event. These systems are designed to capture every move that is made by an individual, though the data are generally assembled into discrete length periods prior to release, or designed to capture changes of residence by extracting information at specified annual intervals. An example of this is Statistics Sweden (SCB, Statistiska Central Byrån) which compiles data collected from various public agencies. The migration registers are created by combining the Total Population Register (Rikets Totalbefolkning – the country's

total population) with the Property-Tax Register (Fastighets Taxerings Registret) with matched information downloaded annually on the last day of December. This means it is possible to follow individuals and their moves on a yearly basis.

In the UK, NHS Digital (formerly the Health and Social Care Information Centre) maintains a demographic database of all registered patients, the Patient Register Data Service (PRDS). A version frozen at the end of each July is used in conjunction with a similar version from one year earlier to generate records of patients who have changed address (recorded as postcodes). These postcode-to-postcode records, classified by age and gender, are used to produce counts of migrants between local authorities; that is, transition data. These transition counts are converted to movement counts by applying ratios of moves to transitions available from a legacy database that counts patient re-registrations for current health administration areas, the National Health Service Central Register (NHSCR). The local authority to local authority counts are supplied to ONS for use in creating mid-year population estimates (ONS, 2012). The NHSCR was not created as a migration registration system, so it provides an indirect method of deriving estimates of relatively long-distance moves taking place between health areas by capturing patient re-registrations with their doctors (Champion and Shuttleworth, 2016a). Although data on all postcode to postcode changes are captured from the PRDS, shorter distance relocations (transitions) are not released for confidentiality reasons. The only moves not captured are those of addresses within a postcode, which are likely to be very rare, as well as those of patients who fail to inform the NHS of their new address when they move (Barr and Shuttleworth, 2012). Medicare registers which provide migration data from the Australian national health system are subject to similar problems and, like British NHSCR data, fail to capture the movements of military personnel. Indeed, selective coverage is endemic to most types of administrative by-product statistics, including such sources as electoral rolls, which are commonly confined to citizens aged 18 and over. Such issues highlight the complexities involved in using administrative data to measure migration and the need for users to be aware of the procedures adopted to generate the data sets that are released. The creation of a migration data time series is complicated further in the UK because different registers are maintained and different methods of estimation are used by the respective national statistical agencies in Scotland and Northern Ireland (see Lomax et al., 2013, for more detail).

Population censuses, in contrast, generally measure migration as a transition. A good example of this is the UK census. Since 1961, UK population censuses have asked a question on address one year before the census, which yields counts of the number of migrants over the 12-month period. Of course, this approach provides only a single snapshot of migration in a particular year and omits moves earlier in each intercensal decade. The Office for National Statistics Longitudinal Study (ONS-LS) provides a longer-term perspective by linking the locations of a sample of individuals from one census to the next. This, in effect, generates a ten-year transition measure which is dependent, of course, on accurately measuring residential locations in successive censuses (Champion and Shuttleworth, 2016b). The weakness of these transition data is that there will be at least some multiple or return moves over a decade about which we know nothing at the start and end of a decade. In the most extreme case, a person might be located in the same place in 2001 and 2011, but have made multiple moves in the intervening period before returning to their original residence. Nevertheless, these longitudinal data prove very useful in providing an indication of the distance over which migrants travel, suggesting a significant decline in the propensity to move relatively short distances (<10km) in England and Wales over the last four decades, as summarised in Chapter 6 and in Champion and Shuttleworth (2016b).

Bell et al. (2002) note that analytically these distinctions between data types are important. Transition data capture migrants whereas event data capture migrations. Differences between the two types of measures are relatively small over short periods, such as a single year, but increase exponentially as the observation interval lengthens. Even over single years, careful harmonisation is needed since transition data measure age at the end of the migration interval whereas event data capture age at the time of migration (see Bell and Rees, 2006). A comprehensive global inventory of the types of internal migration data collected across the 183 of the 193 member states of the United Nations is provided by Bell et al. (2014), who also assess their comparative strengths and weaknesses in detail.

Temporal Comparability

The theme of temporal comparability is complex. One very obvious problem concerns the interval over which migration is measured. It is not possible to compare a question in one country on previous address five years ago with data from elsewhere on address one year ago. In particular, it is not possible to create comparable rates simply by multiplying the one-year

rate by five because five year rates are affected by return and repeat migration, and there is no simple empirical multiplier to place the estimates on a consistent base. As a result, cross-national comparisons must be made separately for countries that collect one year or five-year data (Bell et al., 2002). Moreover, many countries collect information by referring simply to the 'last move', irrespective of timing, while others measure only 'lifetime migration', comparing place of residence at the census with place of birth (Bell et al., 2015c). In this situation, reliable comparisons between countries, or over time, are largely out of reach.

There are also differences in the temporal depth of various data sources. For example, the Swedish Population Register used in Chapter 9 runs from 1990 to 2014 and other Swedish data from 1900, whilst the annual NHSCR time series for England and Wales used in Chapter 6 starts in 1975. Chapter 6 also uses a set of annual estimates of inter-district migration intensity for the UK running from 2001 to 2013 as well as the ONS-LS data in England and Wales from 1971 to 2011, whereas the United States Population Survey of Income Dynamics (PSID), referred to in Chapter 5, started in 1968 and ran annually until 1997 and every other year thereafter. Clearly, considerable caution must therefore be exercised when collating information on temporal trends given these considerable differences between just three countries, all with well-established population data and statistical systems.

Differences in the timing of population censuses also hinder comparability, because they are not necessarily synchronised with fluctuations in the labour and housing markets which often shape migration (Bell et al., 2002). In these situations, it is important to rely on country-specific knowledge and a combination of sources to get a 'best picture' of a reality that might only be seen with difficulty. These problems are compounded by the sensitivity of internal migration to economic cycles which mean that the start and end point of an analysis cannot be ignored (Champion and Shuttleworth, 2016a).

Data Coverage and Quality

One problem with migration data, especially from censuses, is that migrants are hard to enumerate. These are 'hard-to-count' populations because they are mobile – and hence difficult to tie down in statistics – and they also fall into just those demographic groups (for example, the young, students, and those in private rented and communal accommodation) which are problematic for other reasons. It is likely, therefore, that there is an undercount of

migrants to a greater or lesser extent. Administrative systems that record migration events are particularly susceptible to these problems. The NHSCR is based on health identification numbers and internal migration is measured by de- and re-registrations with doctors. However, we know that some groups (such as the younger, healthier and more mobile) tend to lag in re-registering (or even not register at all), which means some migrants are left in the wrong place as far as the registration system is concerned and some moves are unobserved (Stillwell et al., 1992; Barr and Shuttleworth, 2012). This is also a problem faced by the Swedish population register which may undercount the internal migratory moves of young people who leave the parental home if they refuse to register for certain services. At the same time, groups such as military personnel, overseas visitors and recent immigrants, may be omitted entirely from population registers and administrative systems in certain countries since the criteria for inclusion and for registration vary widely between countries and data collections (see Bell et al., 2015c).

There are other analytical issues with data quality and coverage. These are readily exemplified by the problems experienced when trying to analyse internal migration in the UK that arise from changing questions and definitions used in the census. Questions asked at each census vary through time and some topics, for example education, undergo major changes that reflect fast-moving changes as new qualifications start and others end. New questions are also asked: ethnicity, for instance, was introduced to the England & Wales census in 1991. Furthermore, the population base also changed, with students before 2001 being recorded at their vacation (normally parental home) address but in 2001 and 2011 recorded at their term-time address. All these changes make it difficult to make reliable comparisons of internal migration through time. These problems are not restricted to censuses but can also apply to surveys where changed methodologies can lead to discontinuities, as is the case with the United States PSID, and which mean that care must be taken in interpreting changes from the start of the data series in 1968.

Differences in Spatial Units

Internal migration is an inherently spatial phenomenon. It is therefore bedevilled by the problems that are common to all spatial statistics in which measurement is conditioned by the geographies used to capture, output and represent data. The problem of the MAUP is well known and applies across many fields of study (see, for instance, Openshaw, 1984;

Fotheringham and Wong, 1991; Openshaw and Rao, 1995; Holt et al., 1996; Manley, 2014). The gold standard is to have x,y coordinates for individuals or households as is the case in the Swedish population register. This is geo-referenced to a 100-metre resolution, allowing local and longer-distance moves to be accurately measured. Something similar is possible with the ONS-LS where detailed address information collected by the census permits internal migration/housing moves to be defined at a very fine spatial resolution, although it should be noted that the accuracy of this geo-referencing varies between censuses and care must be taken in considering whether all moves or just some over a certain threshold should be considered (Champion and Shuttleworth, 2016b). It is very important to have this finely-grained data since most address changes occur over short distances and the majority of internal migration is thus relatively local. In the UK, data from a consumer survey by a commercial company known as the Acxiom Research Opinion Poll has provided this level of geographical granularity for migrants in Britain over three years in the mid-2000s (Stillwell and Thomas, 2016).

However, except in a few national cases, the migration analyst must work with migration statistics based on pre-defined geographies such as states, regions, parishes, Länder, or other statistical/administrative areas. These are inconsistent in size and shape between countries and this means that estimates of migration distances based on moves between population or geometrical centroids are not comparable. Moreover, administrative and statistical geographies can and do change within countries, making comparisons through time even within the same state problematic (Champion and Shuttleworth, 2016a). One solution is to compile data on the lowest common denominator, but more analytically sophisticated approaches have also been devised. Building on the work of Courgeau (1973; see also Courgeau et al., 2012), Bell et al. (2015a) have used migration intensities measured at a range of different spatial scales to make estimates of all residential moves which are comparable across countries, circumventing the problems caused by differences in spatial frameworks. Coupled with data from the few countries that collect this information directly in the census, this method provided the basis for robust comparisons of migration intensity across 96 countries representing 80% of the global population (see Chapter 4). In a similar manner, Stillwell et al. (2016) have shown how spatial interaction models can be fitted to inter-zonal migration flows to generate distance decay (beta) parameters that capture the effects of distance on migration in a single index that is largely independent of the spatial scale of which migration was measured. Rees et al. (2016) describe another new index, the

Index of Net Migration Impact, which allows reliable cross-national and temporal comparisons of the extent to which internal migration operate to shape the redistribution of population within countries.

Data Availability

Notwithstanding this progress in the development of analytical techniques, cross-national comparisons are fundamentally constrained by data availability. The IMAGE inventory has catalogued what internal migration data are available and from which sources for all countries across the world, indicating that, of the 193 UN member states, 82% collected data from censuses, 26% from administrative sources, 57% from surveys and 56% from multiple sources (Bell et al., 2015b, Table 1). The inventory also indicates that countries collect transition data based on different observation periods (one-year, five-year, other fixed interval, lifetime, last move) and, of particular importance, the variety of forms in which the data collected are released.

If we consider only aggregate migration, there are some countries whose national statistical agencies release origin-destination matrices of migration flows at a number of spatial scales. This is particularly beneficial when directional migration patterns or migration distances are the focus of comparison between countries. Nevertheless, caution is needed in interpreting the diagonal cells of the matrix in some countries since these may contain either counts of intra-zonal flows, counts of intra-zonal movers and stayers or even flows between zones at a lower level in the geographical hierarchy than that for which the data are released. In some countries, this component of the matrix is missing altogether, preventing a figure for total migration in the country from being derived from the matrix; in other countries, only the marginal totals of zonal in-migration and out-migration are available, especially in the case of flows disaggregated by gender or age group (Bell et al., 2015b).

Another common form in which migration data are made available is as simple counts of total internal movement at various spatial scales, such as between states, between counties, or between municipalities. These are often referred to as migration status data, and are the most regular form in which information on population mobility is reported on national statistical websites. Sample surveys also commonly provide data of this type, often accompanied by details of the characteristics of movers or the reasons for migration. This form of count data

provides no information on the spatial pattern of migration flows, but it does provide a crucial measure of the overall intensity of migration at different spatial scales. A small number of countries also collect information on all changes of address irrespective of spatial scale (see Bell et al., 2015b, 2015c), but even where this is unavailable, migration counts at multiple spatial scales provide the essential building blocks to estimate aggregate migration intensities in a form that is comparable between countries, as mentioned above (see Courgeau et al., 2012; Bell et al., 2015a). By the same logic, such counts provide a basis to estimate the trend in migration intensities through time, as explained below, even where regional and local area boundaries have undergone considerable change.

Not all the data collected by countries using various instruments are published in readily available tables or accessible from online information systems and therefore the task of gathering data for comparative analysis and assembling these data is often less than straightforward. The IMAGE inventory is one attempt to accomplish this: it contains internal migration data of various types and forms extracted from repositories (such as the Integrated Public Use Microdata Series-International (IPUMS), the Centro Latinoamericano y Caribeno de Demografia (CELADE), the EUROSTAT database, or the USAID's Demographic and Health Surveys (DHS)) or supplied by national statistical offices in countries around the world, together with the relevant aggregate populations at risk and the boundaries of the geographical zones at each geographical scale for which spatial migration data are available. Many of these data sets, together with GIS boundary files, are now freely available on GitHub (https://github.com).

Finally, we have to recognise that migration data collected by national agencies using census or survey instruments may go through extensive processing before being released as 'official statistics'. The UK census is a particular case in point, with a range of pre-tabulation and post-tabulation adjustments made to create a set of estimates from the raw statistics that meet the confidentiality requirements required under the current legislation. In the case of England and Wales, adjustment methods have changed from one census to the next, creating further uncertainty over the legitimacy of comparison from one census to the next. One specific example of this is the use of small cell adjustment methods (SCAM) in the 2001 Census to ensure that all flows of one or two individuals in the migration tables were changed probabilistically to values of zero or three, consequently rendering the matrices of flows between small areas such as output areas much less useful and limiting the opportunity for

consistent comparison with flows at this spatial scale in 2011 when the SCAM was not applied. In a similar way, new coding procedures were introduced in the 1996 Australian census which resulted in a major disruption to the five yearly census-based time series of migration that dates back to 1976, in this case resulting in a marked upwards shift in the apparent level of local residential mobility (see Chapter 7).

Pathways to Comparability

As indicated in the previous section, there are real problems which limit the reliability of comparisons of internal migration between countries but temporal analysis within countries confronts similar obstacles. The individual country contributions which comprise the heart of this volume are all faced with the impediments described above and have adopted a number of approaches to solve them, or at least minimise their effects.

Creating Consistent Definitions

Some of the solutions are relatively straightforward, such as coding variables to the lowest common denominator and compiling data on a single consistent geography. This was the approach adopted by Champion and Shuttleworth (2016a) in constructing a time series of aggregate migration flows between health areas and regions in England and Wales from the NHSCR from 1975 to 2011. In a similar manner, the chapter on Australia draws on the Australian Internal Migration Database which was carefully constructed using GIS overlays of the basic building blocks (Statistical Local Areas) to create a consistent geography of 69 functional regions (Temporal Statistical Divisions – TSDs) that are spatially harmonised across seven censuses to produce a time series spanning 35 years (Blake et al., 2000). Stillwell et al. (2000; 2001) used TSDs to create a hierarchical structure built around six types of city regions that allowed robust comparisons to be made with migration flows through a similar spatial system constructed from districts in the UK.

Even where concerted attempts are made to harmonise the data, however, there is little basis on which to compare particular forms of migration from one country to another. The supposed distinction between residential mobility and migration is a particular case in point. As Niedomysl (2011) makes patently clear, there is no obvious or easily defined empirical

cut off between local and long-distance migration, and in most countries the distinction is based simply on readily available data for convenient administrative boundaries, which inevitably differ between countries in size, extent and relevance. Thus, what is defined as local in the USA or Australia, is likely to bear little correspondence to data which are similarly described in Japan or the UK. Stillwell et al. (2016) have proposed creative solutions to the problem of comparing countries with respect to migration distance but for the purposes of this volume particular care is needed in making comparisons in regard to 'local' or 'long-distance' mobility, because these will likely measure quite different things across our sample of countries.

All our sample countries suffer errors and inconsistencies due to undercounts and undercoverage to a greater or lesser degree depending on the type of data they use, and these are spelled out in the relevant chapters. In some cases, too, considerable data manipulation has been needed to create a useable time series. In the case of the UK, for example, a nation state that contains four component home nations (England, Wales, Scotland, Northern Ireland) with three national statistical agencies (NSAs), construction of a consistent time series was particularly problematic. The Office for National Statistics (ONS) collects data from the other two NSAs – National Records of Scotland (NRS) and the Northern Ireland Statistics and Research Agency (NISRA) – which are compiled to produce aggregate mid-year population estimates for local authority districts across the UK together with estimates of the components of change using a common methodological approach (ONS, 2011). There are, however, various availability and consistency problems associated with the internal migration data used in the population estimation process by each NSA.

Modelling Data

Chapter 6 reports the temporal changes in migration propensities within the UK that are evident from a time series of estimates that connect the two census periods, 2000-01 and 2010-11, and which has been assembled from data collected from administrative sources used by the three NSAs. An important distinction is drawn between migrant flows between LADs that occur within the same nations and flows between LADs that cross the boundaries between England & Wales and Scotland and Northern Ireland and are described by Lomax et al. (2013) as internal 'international' flows. Whilst administrative sources provide information about all the flows in the former group with the exception of flows between districts in

Northern Ireland, the latter migration flows are unknown and have to be estimated from data on known marginal totals of migration flows between the countries. A number of methods are available to solve the problem of estimating missing data in origin-destination matrices, including log-linear models, gravity models, spatial interaction models, entropy and information maximization models and Iterative Proportional Fitting (IPF). Different techniques have been compared by Willekens (1980; 1983) and, after modelling a multidimensional dataset using different methods, van Imhoff et al. (1997) conclude that IPF is the most efficient in terms of the time taken to generate a solution. IPF was probably first applied to fit a contingency table using marginal constraints by Deming and Stephan (1940) and a comprehensive history of the methodology is provided by Založnik (2011). Details of the methodology underpinning the estimates used in Chapter 6 are available in Lomax et al. (2013) which also reports strong correlations between IPF-derived estimates and observed annual data for the districts of England and Wales derived from PRDS data.

Selecting Robust Migration Indicators

Another issue in making time series comparisons is selection and application of the most appropriate statistical indicators. Bell et al. (2002) suggested that four discrete domains could be recognised for comparison of migration within countries: intensity, impact, distance and connectivity, and specified a set of indicators in each domain which could be used for cross-national comparison. While each of these domains provides a particular perspective on the nature of migration, it is migration intensity, and to a lesser extent migration impact and migration distance, that are of primary relevance to the focus of the current volume. Each of these can be captured using a number of different indicators but by far the simplest and most basic measure for comparison between countries, or over time, is the Crude Migration Intensity (CMI) computed simply as:

$$CMI = M / P \qquad (1)$$

where M represents the number of migrants or migrations at a particular level of spatial scale (e.g., between states or districts) and P is the population at risk (PAR). Following van Imhof et al. (1997), the term intensity is used to encompass both rates and probabilities (see below). While measurement of migration is a primary concern, care is also needed in selection of the appropriate denominator for computation of the CMI.

Whilst a count of those usually resident in a country is generally available from national censuses, this statistic may refer to different points in time. In comparing migration intensities between Britain and Australia, Rees et al. (2000) clarify that different forms of the PAR are needed for event data and for transition data. Event data are distributed throughout the observation interval and require a midpoint population to generate occurrence-exposure rates. Transition data, on the other hand, capture only those who were alive and in the country at the start and end of the interval. In Australia, the PAR for both one and five-year transition probabilities is readily derived from census migration matrices, since these include non-movers as well as those who moved within and between zones, and following Rees et al. (2000) this is used to compute migration intensities. In contrast, the UK census reports only an end-of-period population for the one-year migration data it collects and no start-of-period population is readily available. Theoretically, since a migration event is, on average, likely to take place halfway through the period, the PAR most appropriate for use in the intensity calculation is the mid-period population, which, for census data, requires estimation. This can be obtained by interpolating between mid-year estimates. One benefit of using NHSCR-based event data in the UK is that the time period is from mid-year to mid-year, and the start and end points coincide with date (30 June) for which the ONS produces mid-year sub-national population estimates. Countries vary in regard to the types of data available to represent the PAR, but provided a consistent approach is used, temporal consistency should be maintained.

In making comparisons between countries, Bell et al. (2015a) argued that the only reliable comparison was in terms of all changes of address, irrespective of spatial scale, since countries differed widely in the number of zones or regions for which the CMI was measured. In some countries, data on all changes of permanent address are available from the periodic census but this is often not the case, and is rarely true for population registers or data from administrative sources such as the NHS, when only flows between LADs and between areas used for administering the NHS are available.

To address this problem, the IMAGE project developed an alternative mechanism to estimate the intensity of migration at all spatial scales (the Aggregate Crude Migration Intensity or ACMI), using available information on flows at a range of different spatial scales. The estimates leverage the linear relationship which is found between crude migration intensities

and the average number of households per zone at each spatial scale (see Courgeau, 1973, Courgeau et al., 2012). Where data are available at say three spatial scales (such as regions, provinces and municipalities), the crude migration intensity, CMI(n) can be plotted against the natural log of the average number of households per zone and a regression line can be defined as:

$$CMI(n) = a + b\,(\ln(H/n)) \qquad (2)$$

where H is the aggregate number of households in the country as a whole, n is the number of zones at which the value of the crude migration intensity (CMI) is recorded, b is the slope of the linear association and a is the intercept. When the number of households is the same as the number of zones, then H/n = 1, so that any move represents a migration from one household (or dwelling) to another. The log of H/n then equals zero, and the equation defaults to the value of the Y-intercept, a, which provides an estimate of all changes of address – the ACMI.

Chapter 4 of this book makes use of this technique to provide time-series estimates of the ACMI for 20 countries around the world, including developing countries as well as the developed nations that are the main focus of this volume. However, the data needed to calculate the ACMI are not widely available. Additionally in Chapter 4, therefore, comparisons of trends in migration intensity are made on a between-area-flow basis for a much larger set of countries, 56 in all, using the CMI calculated for the geographies for which comparable time-series data are available in each country.

Like all crude measures, the CMI is influenced by age composition effects, so Bell et al. (2002) set out a number of more sophisticated measures of intensity including age-standardised rates and migration expectancies. Age standardisation is an important consideration when comparing migration trends over time, since population ageing will lead to a reduction in the CMI even if the underlying propensity to move remains unchanged. This effect is considered explicitly in several of the chapters that follow, since it is one of the fundamental potential causes of the observed decline in migration intensity in those countries undergoing population ageing.

Standardisation might also be applied to allow for other compositional effects, such as shifts in educational attainment, or occupational mix, since increasing proportions of highly mobile educated and professional classes are likely to place upwards pressure on mobility. In most analyses such factors are generally introduced as explanatory rather than as structural variables (see Bell et al., 2015a). As Bernard et al. (2014a; 2016) show, however, compositional effects also affect the age at which migration occurs, because the age profile of migration is shaped by the timing of key transitions in the life course, both in the family sphere (partnership and family formation) and in the economic domain (educational participation and labour force entry). Economic development therefore acts not only to change the overall level of migration intensity in a national population, but also the ages at which migration occurs, and this impacts not only on young adults, but also on the age at retirement (see Sander and Bell, 2016).

Bernard et al. (2014b) demonstrate that age and intensity at the peak are the optimum measures to capture cross-national differences in the age profile of migration, since these measures encompass both the breadth and symmetry of the peak among young adults. They also demonstrate (Bernard and Bell, 2015) that considerable care is needed in the choice of technique for data smoothing, since conventional model migration schedules may obscure subtle but important shifts over time in the age at which peak migration occurs. It follows that careful attention to the choice and computation of migration indicators is needed if shifts over time in the overall level of migration within a country, and their underlying causes, are to be properly understood.

Robust indicators are also needed to measure changes over time in the extent to which migration is generating population redistribution within a country. Migration is ultimately a form of spatial behaviour and movements over longer distances commonly take place in response to regional differentials in economic opportunities. As countries develop and become progressively more urbanised, the extent of redistribution between regions rises, slowly at first, then more rapidly, and finally falls away at more advanced stages of economic development (Rees et al., 2016). We know that the extent of population redistribution arising from internal migration can be captured by the Aggregate Net Migration Rate (ANMR) and this in turn is a product of two distinct components: migration intensity and migration effectiveness (Bell et al., 2002), the latter (captured in the Migration Effectiveness Index or

MEI) simply measuring the extent to which movements in one direction are more or less balanced by movements in the other. Thus:

$$ANMR = CMI \times MEI \quad (3)$$

where CMI is defined in equation (1) and:

$$MEI = N/M \quad (4)$$

where N represents the sum of the net migration balances for all gaining regions and M is the total migration between regions.

If longer-distance migration is indeed declining, as appears to be the case in at least some of our sample countries, then it begs the question as to how this decline is interacting with migration effectiveness, and how this then plays out in terms of changes in the level of population redistribution. The MEI, CMI and ANMR provide the tools to trace this interaction over time in those countries where the requisite data are available, as is the case for Australia (see Chapter 7 in this volume).

Conclusions

Individually and collectively, the sample countries in this volume confront all of the data deficiencies and methodological problems outlined in this chapter, though the specifics vary widely from one country to the next. As argued elsewhere (Bell et al., 2002; 2014), there is a strong case to be made for greater uniformity and harmonisation in the way migration data are collected, analysed and made available. In the case of international migration, significant progress has already been made in this regard (see Bilsborrow et al., 1997; Nowok et al., 2006), but for internal migration the task remains largely ahead for scholars and statisticians.

The individual chapter contributions outline the particular challenges of tracing mobility trends in the country of interest and draw on a range of different data types and analytical methods, as well as covering a variety of temporal frameworks. What they all have in common is the endeavour to identify trends in internal migration over an extended period, to

distinguish the forms of migration that have changed, and to attempt to account for the underlying causes of the observed patterns. To establish the broader context for the particular case studies that follow, the next chapter reviews trends in a broader, global sample of countries that include examples drawn from all continents and reflecting a range of cultural settings and stages of economic development. This enables the case study countries to be compared with each other as well as being set in their wider context, identifying issues and arguments which will be returned to in the final part of the book.

References

Barr, P. and Shuttleworth, I. 2012. Reporting address changes by migrants: The accuracy and timeliness of reports via health card registers. Health and Place, 18, 595-604.

Bell, M., Blake, M., Boyle, P., Duke-Williams, O., Rees, P., Stillwell, J. and Hugo, G. 2002. Cross-national comparison of internal migration: issues and measures. Journal of the Royal Statistical Society A, 165(3), 1435-464.

Bell, M, and Charles-Edwards, E. 2013. Cross-national Comparisons of Internal Migration: An Update of Global Patterns and Trends. Technical Paper 2013/1. New York: Population Division, United Nations Department of Economic and Social Affairs. http://www.un.org/en/development/desa/population/publications/pdf/technical/TP2013-1.pdf

Bell, M., Bernard, A., Ueffing, P. and Charles-Edwards, E. 2014. The IMAGE Repository: A User Guide, QCPR Working Paper 2014/01. Brisbane: Queensland Centre for Population Research, University of Queensland.

Bell, M., Charles-Edwards, E., Ueffing, P., Stillwell, J., Kupiszewski, M. and Kupiszewska, D. 2015a. Internal migration and development: Comparing migration intensities around the world. Population and Development Review, 41(1), 33-58.

Bell, M., Bernard, A., Charles-Edwards, E., Kupiszewska, D., Kupiszewski, M., Stillwell, J., Zhu, Y., Ueffing, P. and Booth, A. 2015b. The IMAGE Inventory: A User Guide. QCPR Working Paper 2015/01. Brisbane**:** Queensland Centre for Population Research, University of Queensland.

Bell, M., Charles-Edwards, E., Kupiszewska, D., Kupiszewski, M., Stillwell, J. and Zhu, Y. 2015c. Internal migration data around the world: Assessing contemporary practice. Population, Space and Place, 21(1), 1-17.

Bell, M. and Rees, P. 2006. Comparing migration in Britain and Australia: harmonisation through use of age-time plans. Environment and Planning A, 38(5), 959-988.

Bell, M. and Ward, G.J. 2000. Comparing permanent migration with temporary mobility. Tourism Geographies, 2(1), 97-107.

Bernard, A. and Bell, M. 2015. Smoothing internal migration age profiles for comparative research. Demographic Research, 32 (33), 915-948.

Bernard, A., Bell, M. and Charles-Edwards, E. 2014a. Explaining cross-national differences in the age profile of internal migration: the role of life-course transitions. Population and Development Review, 40(2), 213-239.

Bernard, A., Bell, M. and Charles-Edwards, E. 2014b. Improved measures for the cross-national comparison of age profiles of internal migration. Population Studies, 68(2), 179-195.

Bernard, A., Bell, M. and Charles-Edwards, E. 2016. Internal migration age patterns and the transition to adulthood: Australia and Great Britain compared. Journal of Population Research, DOI 10.1007/s12546-016-9157-0.

Bilsborrow, R., Hugo, G., Oberai, A.S. and Zlotnik, H. 1997. International Migration Statistics: Guidelines for Improving Data Collection Systems. Geneva: International Labour Organization.

Blake, M., Bell, M. and Rees, P. 2000. Creating a temporally consistent spatial framework for the analysis of interregional migration in Australia. International Journal of Population Geography, 6, 155-74.

Champion, A.G. and Shuttleworth, I. 2016a. Is longer-distance migration slowing? An analysis of the annual record for England and Wales. Population, Space and Place, published online in Wiley Online Library, DOI: 10.1002/psp.2024.

Champion, A.G. and Shuttleworth, I. 2016b. Are people moving address less? An analysis of migration within England and Wales, 1971-2011, by distance of move. Population, Space and Place, published online in Wiley Online Library, DOI: 10.1002/psp.2026.

Collier, D. 1993. The comparative method. In Finifter, A.W. (ed) Political Science: The State of the Discipline. Washington, D.C.: American Political Science Association. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1540884

Courgeau, D. 1973. Migrations et découpage du territoire. Population, 28, 511–536.

Courgeau, D., Muhidin, S. and Bell, M. 2012. Estimating changes of residence for cross-national comparison. Population (English edition), 67(4), 631-652.

Deming, W.E., and Stephan, F.F. 1940. On a least squares adjustment of a sampled frequency table when the expected marginal totals are known. The Annals of Mathematical Statistics, 11(4), 427-444.

Duke-Williams, O. 2009. The geographies of student migration in the UK. Environment and Planning A, 41(8), 1826-48.

Fotheringham, A.S. and Wong, D.W.S. 1991. The modifiable areal unit problem in multivariate statistical analysis. Environment and Planning A, 23, 1025-1044.

Holt, D., Steel, D.G. and Tranmer, M. 1996. Area homogeneity and the modifiable areal unit problem. Geographical Systems, 3, 181-200.

Lijphart, A., 1971. Comparative politics and the comparative method. American Political Science Review, 65(03), 682-693.

Lomax, N., Rees, P., Norman, P. and Stillwell, J. 2013. Subnational migration in the United Kingdom: producing a consistent time series using a combination of available data and estimates. Journal of Population Research, 30, 265-288.

Long, L., Tucker, C.J. and Urton, W.L. 1988. Migration distances: an international comparison. Demography, 25(4), 633-640.

Manley, D. 2014. Scale, aggregation, and the Modifiable Areal Unit Problem. In Fischer, M.M. and Nijkamp, P. (eds) Handbook of Regional Science. Berlin: Springer-Verlag, 1157-1171.

Niedomysl, T. 2011. How migration motivations change over migration distance: evidence on variations across socioeconomic and demographic groups. Regional Studies, 45(6), 843-55.

Niedomysl, T., Ernstson, U. and Fransson, U. 2017. The accuracy of migration distance measures. Population, Space and Place, 23(1), e1971. DOI: 10.1002/psp.1971.

Nowok, B., Kupiszewska, D. and Poulain, M. 2006. Statistics on international migration flows. In Poulain, M., Perrin, N. and Singleton, A. (eds) Towards Harmonised European Statistics on International Migration (THESIM). Louvain-la-Neuve: Presses Universitaires de Louvain, 203-231.

ONS. 2011. Mid-year Estimates Short Methods Guide. Titchfield, UK: Office for National Statistics. http://www.ons.gov.uk/ons/guide-method/method-quality/specific/population-and-migration/pop-ests/mid-year-population-estimates-short-methods-guide.pdf. Accessed August 14, 2012.

ONS. 2012. A Conceptual Framework for UK Population and Migration Statistics. Titchfield, UK: Office for National Statistics.

http://webarchive.nationalarchives.gov.uk/20160105160709/http://www.ons.gov.uk/ons/guide-method/method-quality/imps/latest-news/conceptual-framework/index.html

Openshaw, S. 1984. The Modifiable Areal Unit Problem. CATMOG 38. Norwich: GeoBooks.

Openshaw, S. and Rao, L.1995. Algorithms for re-engineering 1991 census geography, Environment and Planning A, 27, 425-446.

Rees, P., Bell, M., Blake, M. and Duke-Williams, O. 2000. Harmonising databases for the cross national study of internal migration: lessons from Australia and Britain. Working Paper 00-05. Leeds: School of Geography, University of Leeds.

Rees, P., Bell, M., Kupiszewski, M., Kupiszewska, D., Ueffing, P., Bernard, A., Charles-Edwards, E. and Stillwell, J. 2016. The impact of internal migration on population redistribution: an international comparison. Population, Space and Place, published online in Wiley Online Library, DOI: 10.1002/psp.2036.

Sander, N. and Bell, M. 2016. Age, period and cohort effects of migration on the baby-boomers in Australia. Population, Space and Place, 22(6), 617-630.

Stillwell, J., Bell, M., Blake, M., Duke-Williams, O. and Rees, P. 2000. A comparison of net migration flows and migration effectiveness in Australia and Britain, 1976-96. Part 1: Total migration patterns. Journal of Population Research, 17(1), 17-41.

Stillwell, J., Bell, M., Blake, M., Duke-Williams, O. and Rees, P. 2001. A comparison of net migration flows and migration effectiveness in Australia and Britain, 1976-96: Part 2, Age-related migration patterns. Journal of Population Research, 18(1), 19-39.

Stillwell J., Bell M., Ueffing P., Daras K., Charles-Edwards E., Kupiszewski M. and Kupiszewska, D. 2016. Internal migration around the world: comparing distance travelled and its frictional effect. Environment and Planning A, 48(8), 1657-1675.

Stillwell, J., Rees, P.H. and Boden, P. (eds) 1992. Migration Processes and Patterns Volume 2: Population Redistribution in the United Kingdom. London, Belhaven Press.

Stillwell, J. and Thomas, M. 2016. How far to internal migrants really move? Demonstrating a new method for the estimation of intra-zonal distance. Regional Studies, Regional Science, 3, 28-47.

van Imhoff, E., van der Gaag, N., van Wissen, L. and Rees, P. 1997. The selection of internal migration models for European regions. International Journal of Population Geography, 3(2), 137-159.

Willekens, F.J. 1980. Entropy, multiproportional adjustment and the analysis of contingency tables. Systemi Urbani, 2(3), 171-201.

Willekens, F.J. 1983. Log-linear modelling of spatial interaction. Papers of the Regional Science Association, 52, 187-205.

Založnik, M. 2011. Iterative proportional fitting: Theoretical synthesis and practical limitations. Unpublished PhD thesis. Liverpool: University of Liverpool.